AD-A100 966 SCHOOL OF AEROSPACE MEDICINE BROOKS AFB TX F/G 5/10
EVALUATION OF A MANIKIN PSYCHOMOTOR TASK.(U)
MAY 81 D C READER, R A BENEL, A J RAHE
UNCLASSIFIED SAM-TR-81-10 NL

1 OF 1
AD A
100966

END
DATE
FILMED
7-81
DTIC

LEVEL II

Report SAM-TR-81-10

EVALUATION OF A MANIKIN PSYCHOMOTOR TASK

David C. Reader, Wing Commander, RAF
Russell A. Benel, Ph.D.
Alton J. Rahe, M.S.

May 1981

Final Report for Period January 1980 - October 1980

USAF SCHOOL OF AEROSPACE MEDICINE
Aerospace Medical Division (AFSC)
Brooks Air Force Base, Texas 78235

NOTICES

This final report was submitted by personnel of the Crew Performance Branch, Crew Technology Division, and Data Sciences Division, USAF School of Aerospace Medicine, Aerospace Medical Division, AFSC, Brooks Air Force Base, Texas, under job order 7930-10-37.

When U.S. Government drawings, specifications, or other data are used for any purpose other than a definitely related Government procurement operation, the Government thereby incurs no responsibility nor any obligation whatsoever; and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data is not to be regarded by implication or otherwise, as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

The voluntary informed consent of the subjects used in this research was obtained in accordance with AFR 169-3.

This report has been reviewed by the Office of Public Affairs (PA) and is releasable to the National Technical Information Service (NTIS). At NTIS, it will be available to the general public, including foreign nations.

This technical report has been reviewed and is approved for publication.

DAVID C. READER, Wing Commander, RAF
Project Scientist

BRYCE O. HARTMAN, Ph.D.
Supervisor

ROY L. DEHART
Colonel, USAF, MC
Commander

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: 14 SAM-TR-81-10
2. GOVT ACCESSION NO: AD-A100 966
3. RECIPIENT'S CATALOG NUMBER:
4. TITLE (and Subtitle): 6 EVALUATION OF A MANIKIN PSYCHOMOTOR TASK.
5. TYPE OF REPORT & PERIOD COVERED: 9 Final Rep[redacted]t. Jan[redacted]-Oct[redacted]80
6. PERFORMING ORG. REPORT NUMBER:
7. AUTHOR(s): 10 David C. Reader, Wg Cdr, RAF; Russell A. Benel, Ph.D.; Alton J. Rahe, M.S.
8. CONTRACT OR GRANT NUMBER(s):
9. PERFORMING ORGANIZATION NAME AND ADDRESS: USAF School of Aerospace Medicine (VNE), Aerospace Medical Division (AFSC), Brooks Air Force Base, Texas 78235
10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 62202F 16 7930-10-37 17 1!
11. CONTROLLING OFFICE NAME AND ADDRESS: USAF School of Aerospace Medicine (VNE), Aerospace Medical Division (AFSC), Brooks Air Force Base, Texas 78235 11
12. REPORT DATE: May [redacted]81 12
13. NUMBER OF PAGES: 23 27
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office):
15. SECURITY CLASS. (of this report): Unclassified
15a. DECLASSIFICATION DOWNGRADING SCHEDULE:
16. DISTRIBUTION STATEMENT (of this Report): Approved for public release; distribution unlimited.
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report):
18. SUPPLEMENTARY NOTES:
19. KEY WORDS (Continue on reverse side if necessary and identify by block number):

Manikin Task
Psychomotor Tasks and Performance
Plateau Performance, Objective and Subjective Performance Measures
Task Differential Stability

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

A small, self-contained microprocessor displays a manikin on a TV screen in various orientations. Subjects are required to choose lateralities depending on a shape also presented on the display, and performance (accuracy and reaction time) is automatically recorded.

This task has been evaluated at the USAF School of Aerospace Medicine. Data are presented to show that the task is simple to use, can be learned by all subjects, and that plateau performance can be quickly established.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. ABSTRACT (Continued)

The various presentations were examined for difficulty; subjective and objective measures of performance were compared; and repetitive effects were sought with similarities of presentation in sequence.

The analyses showed that control performance with this task was essentially constant over the experimental period.

CONTENTS

FIGURES

TABLES

CONTENTS (Continued)

EVALUATION OF A MANIKIN PSYCHOMOTOR TASK

INTRODUCTION

Reaction time as a measure of human performance has been in use for over 100 years. Response to a single stimulus is referred to as simple reaction time, while the multiple stimulus situation is known as choice reaction time. Preference for reaction time as a measure of human performance has been predicated on the proposition that behavior takes time and the length of time recorded as reaction time is an accurate appraisal of at least some segment of human behavior. Since reaction time is commonly defined as the time from the onset of a stimulus to the initiation of the subject's response, cognitive rather than motor processes are reflected in reaction time. To exploit this measurement fully, more complex stimulus arrays are necessary. For as William James (3) noted, it is only when the experiments are complicated that there is a chance for anything like an intellectual operation to occur.

Successful performance in modern military systems is, to a large extent, less dependent on skillful physical manipulation of controls, than it is on careful planning of resource management and supervision of automated or semi-automated system operation. At the same time highly trained and skilled operators are typically faced with complex stimulus arrays which nonetheless require discrete control actions. An experimenter who seeks to tap similar psychomotor behavior in the laboratory with unskilled or minimally trained subjects faces a difficult decision in task selection. For these reasons more cognitively based performance measures that are not system dependent have become increasingly important for laboratory performance measurement. The manikin task was developed at the Royal Air Force (RAF) Institute of Aviation Medicine, Farnborough, United Kingdom, as a complex reaction time task that related to performance in modern systems without being identified with any particular system.

For the manikin task (or any task) to be useful as a laboratory analog of some component of real-world system performance, it must meet several practical criteria. Since time is generally limited for training of subjects, acquisition of asymptotic performance should not require protracted training. Also, it is impractical for all subjects to receive training on precisely the same schedule. Therefore, asymptotic performance should not be training-schedule dependent. If preasymptotic performance is to be useful as a measure, then the relative positions of individuals' performance should be constant across some period of training and differential task stability (reliability) should be evident. For the task to have validity, it should be shown to be sensitive to manipulations of interest; e.g., the manikin test has been shown to be sensitive to mild hypoxia.

Although the validity of the manikin task appeared well-established and had been used to measure performance during hypoxia (2), reliability had not been systematically investigated. Additional factors influencing acquisition of asymptotic performance also remained unidentified. For these reasons data collection was instituted to verify the utility of the manikin task. Additional analyses were undertaken to determine intratask variables that related to subject performance. Subjective impressions of performance were compared with actual measures of performance in a later phase of the experiment. These data were also analyzed to seek repetition effects.

Format of Task

The apparatus consists of a microprocessor, power supply box, TV monitor, and two hand-held buttons (Fig. 1). A digital recorder is also required.

The manikin task presents a video picture of a little man (manikin) in outline (Fig. 2). The arms are outstretched and hold a square in the right hand and a disc in the left hand, or vice versa. The manikin may be either upright or inverted and, in addition, may face towards or away from the subject (four types of presentations). Facial and clothing details are visible in the facing presentation. Below the manikin either a disc or square is shown. The subjects' task is to determine in which hand the manikin holds the identical shape to that presented below it and, having decided, to press the button corresponding to that hand.

The manikin picture is generated in a small microcomputer which also controls the presentation timing. Each manikin is presented for 2 seconds and occluded for 1 second. In addition, the manikin is occluded following each response. A series of 96 manikins is presented in sequence. The sequence is fixed, but it always starts at a different place. The 96 presentations consist of 6 segments of 16 presentations, and each segment contains all the 16 possibilities of presentation but in a different random order. Thus, for the purpose of analysis, each segment of 16 (48 seconds) is of equal difficulty. The 96 manikin presentations make a period of 4.8 minutes. Four periods interspersed with 3 rest periods of 2 minutes are grouped to make a session of 25.2 minutes. The experimenter presses the "prime" button on the microcomputer to start a session of testing. Before each period the subject is warned by a tone for 4 seconds before the first manikin presentation of that period. At the end of the last period, a warbling tone informs that the session is ended and that the microprocessor timer requires restarting by pressing the "reset" button. Pressing the reset button before the end of the session resets the timer, and the prime button restarts it.

The microcomputer also generates a recording signal of the psychomotor data. When linked with a digital recorder, the following data are recorded after each presentation: a number from 0-16 denoting the type of presentation; after a space, a 1, 0, or * denoting whether a correct, incorrect, or no response was recorded; and after a further space, the reaction time in seconds to 2 decimal places. At the end of the session, a summary statement of total numbers of correct, incorrect, and no responses and the aggregate of reaction time is recorded.

The logic of the microprocessor is fixed; neither the timing nor presentation sequence can be altered.

The manikin task data were recorded on the disc storage of a PDP 11/34 computer. An ASR 40 line printer was used in parallel to the computer to record a hard copy of the data. A small switch box was employed whereby indices containing subject and run details could be inserted into the disc storage by the line printer.

Experimental Protocol

Two experiments were performed using the manikin task. The first sought to examine the variables that govern acquisition of plateau performance, i.e., those that control the shape of the learning curve. Eighteen naive subjects were exposed to 10 sessions of training. They were divided into the following groupings: (a) by training frequency--6 subjects trained twice a day, 6 subjects trained once a day, and 6 subjects trained every alternate working day; (b) by age--below 28 years (7), between 28 and 38 years (7), and over 38 years (4); and (c) by occupational level--viz, professional scientists (6), noncommissioned officers (6), and airmen (6). Subjects were not permitted to examine the data during the study.

The second experiment examined the relationship of subjective and objective measures of performance. Subjective assessment of performance has been used extensively where direct objective measures of performance are denied. The same 18 subjects were exposed to five further sessions with the manikin task, one per day, on consecutive days. They were asked to estimate both the workload effort and their performance against seven-point scales (Figs. 3 and 4). This experiment also examined whether the shape (disc or square) or side (right or left) affected overall performance and also checked the difficulty of each of the four types of presentations for the subjects.

Repetition effects were also considered using data from the second experiment. The manikin task sequence is fixed; thus all subjects were presented with identical sequences (but different start positions). The similarity of presentations to the immediate preceding presentation varies as shapes, sides, upright or inverted, forward and away facing, are altered. The sequences permitted 7 levels of similarity to be compared from exactly comparable to gross dissimilarity (every variable reversed). The data from the sessions conducted during the second experiment (subjective versus objective assessment) were analyzed to detect repetition effects, i.e., enhanced performance with similarity of presentations in sequence. Finally, an analysis was performed on the stability or reliability of the task data over the entire period of the experiment.

Experimental Procedure

Subjects sat in a darkened, quiet room, viewed a 7-inch (17.78 cm) video display from approximately 38 inches (96.52 cm) (approximately 10° visual angle) and held press-button boxes in either hand. Only one session of performance (25.2 min) was conducted for each subject at a time. In some instances, ear defenders were used when external noise could not be controlled. An experimental monitor briefed the subjects and controlled the apparatus and computer. Distraction to the subjects was minimized at all times.

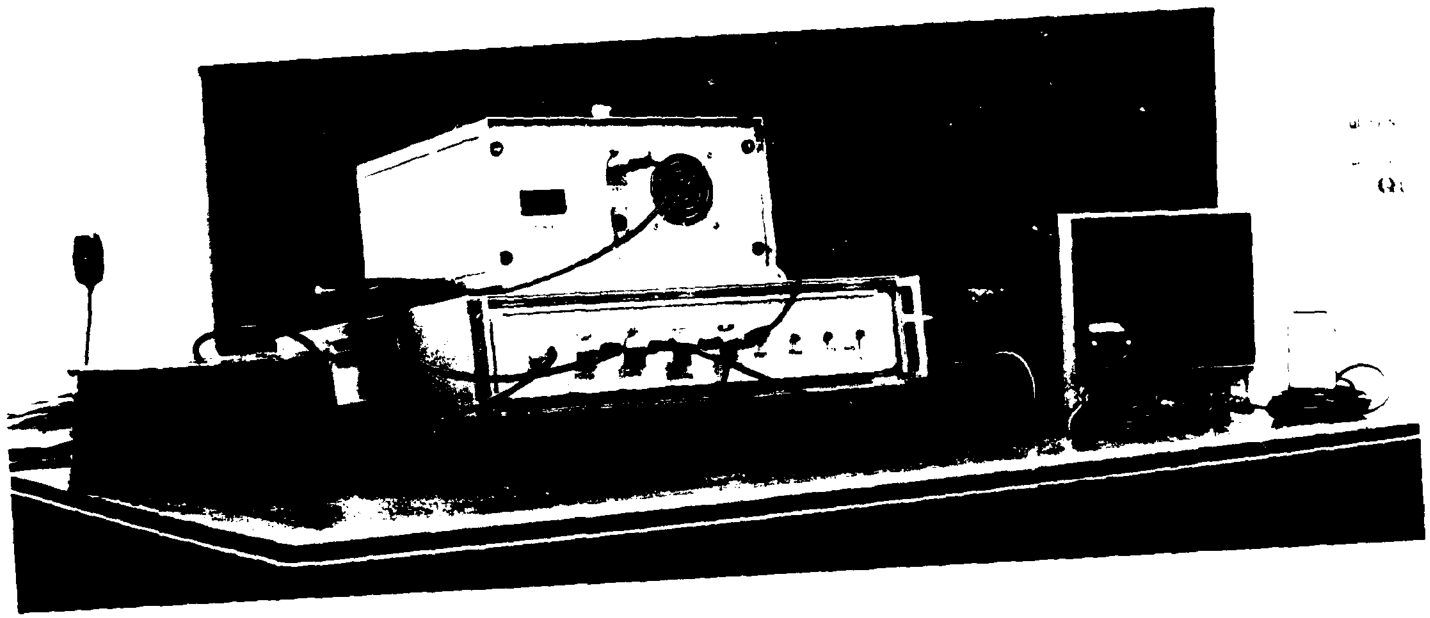
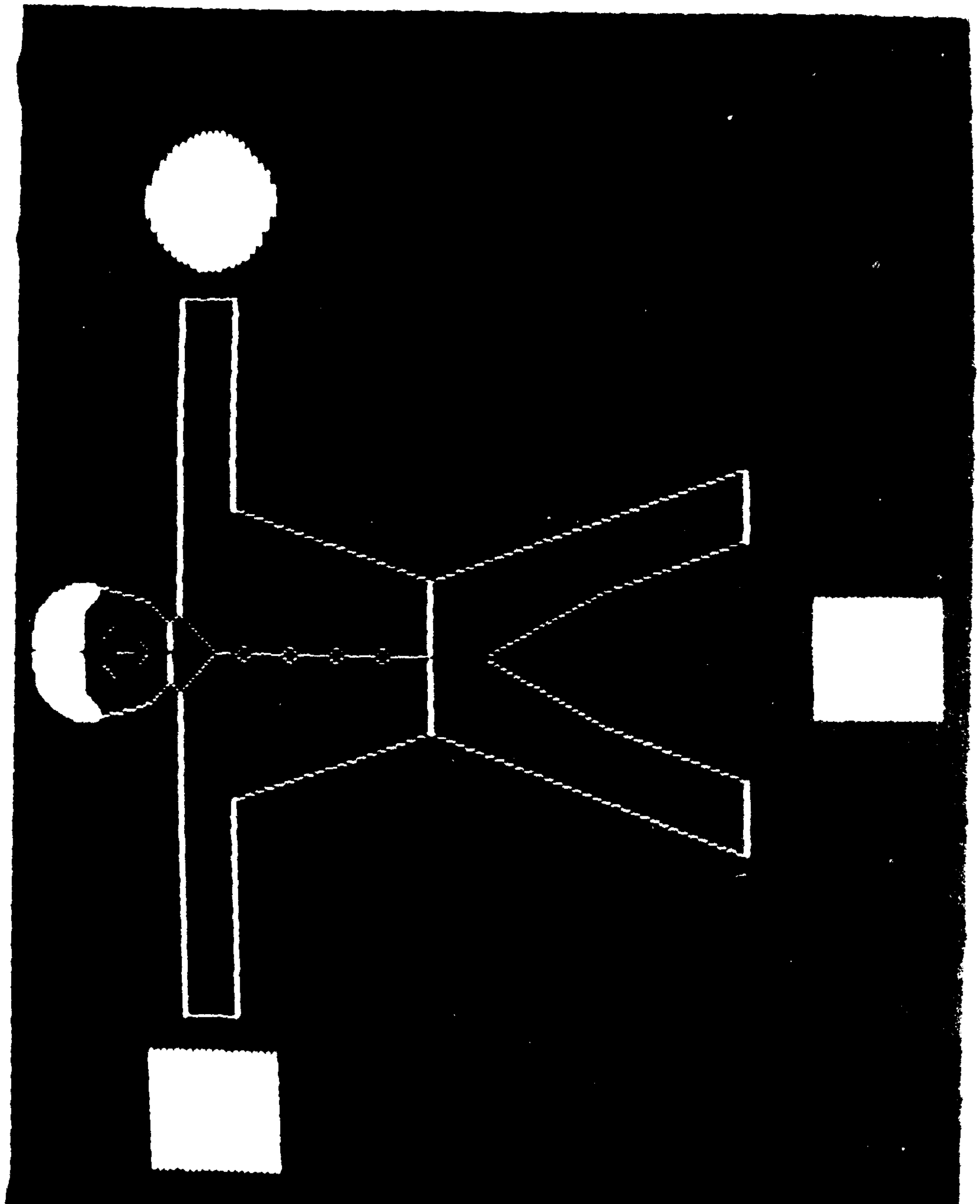


Figure 1. The manikin apparatus.



INSTRUCTIONS: Circle the number that best describes the workload you experienced while performing the task during this session.

1. NOTHING TO DO; NO SYSTEM DEMANDS
2. LITTLE TO DO; MINIMUM SYSTEM DEMANDS
3. ACTIVE INVOLVEMENT REQUIRED: EASY TO KEEP UP
4. CHALLENGING, BUT MANAGEABLE
5. EXTREMELY BUSY; BARELY ABLE TO KEEP UP
6. OVERLOADED; HIGH CHANCE OF ERROR
7. UNMANAGEABLE; UNACCEPTABLE

Figure 3. Workload estimate.

INSTRUCTIONS: Circle the number of the appropriate term. Compared to my performance last session, my performance on this session was...........

1. MUCH WORSE
2. WORSE
3. SLIGHTLY WORSE
4. SAME
5. SLIGHTLY BETTER
6. BETTER
7. MUCH BETTER

Figure 4. Performance estimate.

The data on disc storage was transferred to an IBM 360 computer where all analyses were performed using the Statistical Analysis System (SAS).

RESULTS

Experiment 1 examined the acquisition of plateau performance. For the purpose of the experiment, plateau performance was defined as not exceeding ± 5% of the mean reaction time (RT) of the previous 2 sessions. The number of sessions required to achieve this criterion are given in Table 1. The data are classified by occupational level. In one case, a strategy adopted by one subject (JB) allowed him to improve his performance throughout the experiment. Therefore, he did not reach the criterion set and was excluded from the analysis. Table 2 classifies the data by frequency of training. Table 3 classifies the number of sessions to plateau by age. Both a one-way analysis of variance and a nonparametric H test on ranked data failed to detect any significant differences in number of sessions to plateau between any of the above three groupings (training, occupational level, and age) at the 0.1 level.

A repeated measurements analysis of variance procedure was used to analyze the reaction time (RT) data for training group, session, and period differences, plus their interactions. (See Table 4 for the sources of variation in the analysis.) There was strong evidence ($p<.001$) that the pattern of the period means was not consistent from one session to another (period by session interaction). This inconsistency is a reflection of more improvement over periods for the initial sessions than for the latter sessions, a finding which was expected. The analysis failed to detect any differences among the three training groups in this improvement in performance. That is, the training group did not show any significant interaction with period or session. Of course, the overall performance improved over the ten sessions ($p<.001$).

There was some indication of an overall difference among the three training groups ($p<.10$). (See Table 5 for means.) The group that trained on alternate days had a lower mean reaction time than that of the other two groups. The repeated measurements analysis procedure was also used to analyze the "number of correct responses" and the mean reaction time for the correct responses. The testing results were very comparable among the analyses for the three measurements. Again the group that trained on alternate days had a better accuracy mean (more correct responses) and lower mean reaction time for the correct responses only. It must be remembered that these are overall training group mean differences (averaged over all sessions) and therefore do not reflect a training effect.

The associated mean reaction time, etc., are also given in Table 5 for the age and occupational level groupings. Generally, no significant differences were detected among these groupings. The analyses on each of the three groupings (training, occupational level, and age) were performed ignoring the other two groupings.

Further analyses of the data showed that the manikin presentations posed varying degrees of difficulty. Table 6 lists the mean reaction time for the three groups divided by presentation type.

TABLE 1. NUMBER OF SESSIONS TO CRITERION (PLATEAU PERFORMANCE) (CLASSIFIED BY OCCUPATIONAL LEVEL)

Group	Subject	Number	Mean	Min	Max
Professional scientist	WS	4			
	SG	4			
	DB	10			
	LP	5			
	PD	8			
	DM	5			
			6	4	10
NCOs	BW	5			
	PP	7			
	EC	7			
	FA	5			
	DT	5			
	JB	* (Did not reach plateau)			
			5.8	5	7
Airmen	RC	6			
	TB	7			
	VP	9			
	TW	5			
	JS	5			
	MF	7			
			6.5	5	9

TABLE 2. NUMBER OF SESSIONS TO CRITERION (PLATEAU PERFORMANCE) (CLASSIFIED BY FREQUENCY OF TRAINING)

Group	Subject	Number	Mean	Min	Max
Twice daily	PD	8			
	DM	5			
	FA	5			
	DT	5			
	JS	5			
	MF	7			
			5.8	5	8
Once daily	DB	10			
	LP	5			
	PP	7			
	EC	7			
	VP	9			
	TW	5			
			7.1	5	10
Alternate days	WS	4			
	SG	4			
	BW	5			
	JB	* (Did not reach plateau)			
	RC	6			
	TB	7			
			5.2	4	7

TABLE 3. NUMBER OF SESSIONS TO CRITERION (PLATEAU PERFORMANCE) (CLASSIFIED BY AGE)

Group	Subject	Number	Mean	Min	Max
>38	WS	4			
	DB	10			
	PD	8			
	PP	7			
			7.25	4	10
28-38	SG	4			
	LP	5			
	DM	5			
	FA	5			
	EC	7			
	BW	5			
	JS	5			
			5.1	4	7
<28	JB	* (Did not reach plateau)			
	DT	5			
	RC	6			
	TB	7			
	VP	9			
	TW	5			
	MF	7			
			6.5	5	9

TABLE 4. SOURCES OF VARIATION FOR THE REPEATED MEASUREMENTS ANALYSIS

Source	Degrees of freedom
Between training groups	2
Between subjects* within training groups	14
Between sessions	9
Interaction of training groups and sessions	18
Interaction of subjects and sessions within groups	134
Between periods	3
Interaction of training groups and periods	6
Interaction of subjects and periods within groups	44
Interaction of sessions and periods	27
Interaction of training groups and sessions and periods	54
Interaction of subjects and sessions and periods within groups	404

*Data for JB were not included in the analysis.

TABLE 5. PERFORMANCE IN THE THREE GROUPS

Group	ΣRT (sec)	No. correct	Mean RT (sec) for correct responses
Training frequency			
Twice daily	78.18	90.8	.809
Once daily	88.22	91.4	.912
Alternate days	68.24	93.1	.709
Occupational level			
Professional scientist	83.38	91.5	.865
NCOs	79.37	91.1	.820
Airmen	71.08	92.7	.738
Age			
<28	70.57	92.6	.732
28-38	79.51	92.1	.822
>38	87.15	89.9	.901

TABLE 6. PERFORMANCE AND PRESENTATION TYPE

Group	Mean reaction time (sec) for correct responses			
	Erect away	Erect facing	Inverted facing	Inverted away
Training frequency				
Twice daily	.741	.803	.830	.864
Once daily	.845	.921	.929	.956
Alternate days	.643	.728	.705	.762
Total	.793	.817	.821	.860
Occupational level				
Professional scientist	.799	.878	.869	.907
NCOs	.729	.798	.866	.896
Airmen	.694	.769	.720	.772
Age				
<28	.684	.750	.727	.771
28-38	.740	.815	.846	.892
>38	.826	.913	.923	.947

A repeated measurements analysis was performed for each presentation type separately using only the RT values for which the associated responses were correct in order to eliminate the reaction times with the incorrect responses. The results for each analysis were very comparable to the previous results; i.e., no additional meaningful information was gained. Generally, the erect away presentation had the lowest mean reaction time with the erect facing, inverted facing, and inverted away groups having progressively increasing reaction times. The means are also given in Table 6 for the occupational level and age groupings. No statistical tests were performed on these groupings.

The second experiment examined the correlation of subjective and objective measures of performance. After each session the subjects reported on a scale from 1-7 their impressions of workload and performance. These data were compared with the performance data (RT) from the manikin task. As subject JB used a different strategy for the task than all the other subjects, his data were omitted from the analysis. Table 7 lists the Pearson product-moment correlation coefficients between the mean RT data (for all sessions) and each of these two subjective values among the 17 subjects. These values are low and do not differ statistically from zero. Table 8 lists the correlation coefficient for performance and RT as well as workload and RT among the 5 sessions within each subject separately. Only one (.898) of the coefficients differed significantly from zero. However, each coefficient is based on only 5 pairs of data. The pooled correlation coefficients do not differ significantly from zero.

The reaction time data for the second study were also used to test for shape (square vs. circle), side (right vs. left), and type of presentation (4) differences. A repeated measurement (shape, side, type, and session) analysis of variance procedure was used to test for differences in reaction time. Session did not show any significant interaction with any of the other three factors. Table 9 lists the mean reaction time for the various types of manikin presentations. The pattern of the RT means for shape and side are generally quite comparable for the different presentation levels. The types of presentation were ranked erect away, inverted facing, erect facing, and inverted away in ascending order of reaction time and difficulty ($p=.001$). Table 9 also shows the effect of shape and side. Discs did not differ from squares, but there was a significant ($p<.001$) difference of right from left. The right-sided responses were approximately 25 msec faster than the left.

Performance with emphasis on the similarity of successive presentations was also studied. All presentations were classified by degree of similarity from identical (1) to gross dissimilarity (7) as described in Table 10. The mean reaction times for the second presentations of each successive pair for these sequences are listed in Table 10. The order of similarity is not echoed exactly in the mean reaction time ranking, but a similar trend is seen. The reaction times with the grossly dissimilar sequences (similarity 6 and 7) differ significantly ($p<.01$) from the moderately dissimilar sequence reaction times. In order to equalize the variances in this analysis, a log transform of each RT value was used in a two-way (proportional) analysis of variance.

TABLE 7. CORRELATION BETWEEN OBJECTIVE AND SUBJECTIVE PERFORMANCE RATINGS

	RT vs. workload	RT vs. performance
Mean RT across all sessions (N=17)	.028	-.337

TABLE 8. SUBJECTIVE VS. OBJECTIVE ASSESSMENTS

Subjects	Reaction time compared with Performance	Reaction time compared with Workload	Performance Min	Performance Max	Workload Min	Workload Max
WS	.048	-.714	4	5	3	4
SG	-.595	.786	3	6	2	4
DB	-.838	.726	3	5	1	2
LP	-.189	.758	3	6	2	3
PD	-.100	.000	3	6	4	4
DM	-.488	-.253	3	6	2	3
BW	-.075	.000	4	5	2	2
PP	-.286	.323	3	5	2	3
EC	.000	.000	4	4	2	2
FA	-.434	.463	3	5	2	3
DT	-.178	.000	3	5	2	2
JB	Excluded					
RC	-.594	.787	3	5	3	4
TB	-.032	.898	2	7	2	4
VP	.311	-.809	3	5	2	3
TW	-.189	.492	1	4	2	3
JS	-.221	.000	3	7	3	3
MF	.270	.000	4	6	2	2
Pooled	-.244	.264				

TABLE 9. MEAN REACTION TIME FOR VARIOUS PRESENTATIONS

Shape	Side	Presentation: Erect away	Erect facing	Inverted facing	Inverted away	Total
Circle	Left	.626	.664	.670	.713	.668
	Right	.610	.651	.637	.678	.644
Square	Left	.630	.670	.658	.705	.666
	Right	.612	.635	.648	.664	.640
	Total	.620	.655	.653	.690	

TABLE 10. MEAN REACTION TIME WITH VARYING DEGREES OF SIMILARITY IN SEQUENCE

Similarity	Description	N*	Mean RT	Order
1	Same presentation, shape, side	19	.615	1
2	Same presentation, side	39	.636	5
3	Same presentation	156	.619	4
4	All erect, same side All inverted, same side	373	.618	3
5	All erect All inverted	314	.616	2
6	All different but same side	471	.644	6
7	All different	510	.645	7

3, 4, 5 differ significantly from 6, 7. ($p<.01$)

*Number of tasks/subjects.

A major criterion of any psychomotor task is that it should produce stable results in stable conditions; i.e., the differential stability should be high. Bittner (1) has suggested that all tasks should be examined for stability and proposed a graphic analysis strategy for this. Pearson product-moment correlation coefficients were calculated between all pair-wise session means (1-15) using the reaction data with only correct responses for the 17 subjects. These are listed in Table 11. The lowest between-sessions correlation was .56 (sessions 1-13), indicating a fair degree of stability even without extensive practice. A selection of these correlations is plotted on Figure 5 for sessions 1, 2, 4, 6, 10, and 12. Moreover, the manikin task may be seen to have good task reliability with an estimate of the common correlation of .84.

TABLE 11. CORRELATIONS OF MEAN PERFORMANCE WITH TIME

Session	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	.82	.73	.70	.77	.71	.68	.60	.74	.65	.62	.59	.56	.65	.66
2		.94	.94	.95	.89	.88	.80	.87	.90	.77	.77	.73	.77	.82
3			.98	.98	.94	.89	.86	.87	.92	.69	.72	.67	.72	.85
4				.98	.95	.91	.87	.88	.93	.73	.77	.70	.75	.84
5					.95	.91	.87	.92	.92	.72	.76	.71	.74	.87
6						.94	.91	.91	.94	.71	.75	.69	.75	.82
7							.97	.97	.97	.85	.85	.78	.84	.86
8								.95	.95	.80	.81	.76	.80	.87
9									.93	.80	.81	.74	.79	.85
10										.84	.87	.82	.85	.89
11											.97	.92	.94	.81
12												.93	.92	.84
13													.95	.90
14														.88
15														

DISCUSSION

The choice of a criterion for plateau performance was based on acceptable tolerance levels for control sessions for individual subjects. That all but one subject achieved plateau performance in the confines of the experimental exposures was gratifying. The strategy adopted by one subject (JB) was interesting and unique. He reasoned that the logic of the manikin microprocessor demanded a fixed sequence, albeit with a random start position. Having decided that, he sought to validate his hypothesis by memorizing sequences and hunting those sequences in later exposures. In that event he was very successful. He demonstrated memory scans of up to 30 sequences where the reaction time was zero and correct responses were recorded. Accordingly, as his memory bank grew, he achieved better performance and never achieved a plateau. As he was not dissuaded from that strategy, he persisted; thus his data are misleading and were excluded from the analysis.

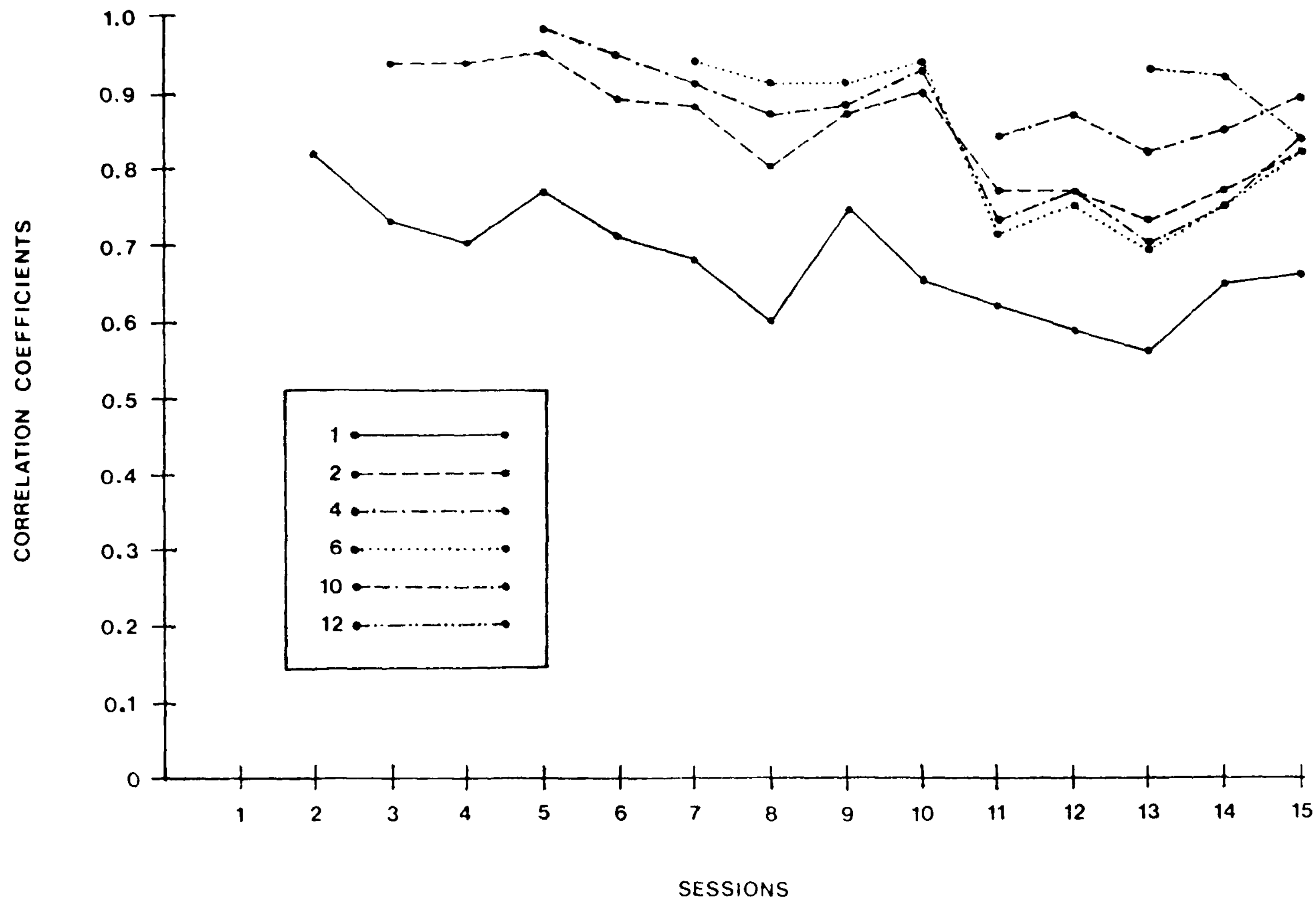


Figure 1. Task differential stability.

No subject achieved plateau performance with less than 4 sessions of training, but all had reached the criterion in 10. The mean was 5.06, and this figure gives good guidance for future training sessions. The data failed to suggest that frequency of training, age, or occupational level affect the rate of achievement of plateau performance which enhances the utility of the task. Admittedly, the age span used in the experiment was small, but it was the maximum practical, given the available subjects.

Although the number of training sessions required for plateau performance showed no appreciable differences with the groups used, the overall performance did provide some suggestions that the group that trained on alternate days had both a lower mean reaction time and a better record of accuracy. However, the number of subjects in the training groups was small (6) and these observed, albeit significant, differences were due more to individual differences than predictable group distinctions. Furthermore, the groupings of age and occupational level were not so distinguished, and these data were recorded before plateau performance had been achieved.

The analysis showed that all subjects considered the types of manikin presentation as the following order of difficulty as graded by the mean reaction times for each: erect away, erect facing, inverted facing, and inverted away. This corresponds to previous analyses which led to the adoption in the manikin software of the even distribution of all possibilities of presentation in each 16 presentations. This permits analysis of performance for each 16 presentation segment, if need be; e.g., for rapid changes in performance or consciousness. However, no such segmental analyses were performed in this study. The order of difficulty is a reflection of the complexity of the mental reorientation required to match the manikin to that of the subject. The upright stance was clearly easier than the inverted, the facing away better than facing towards. The order of difficulty was preserved across the frequency of training groups, with four minor exceptions.

In the second experiment a similar analysis showed a very small variation in this sequence; i.e., the exchange in the ranking of erect facing and inverted facing. Experiment 1 data were from subjects who were learning the task; those from Experiment 2 came from trained personnel. The difference in the order may indicate some changes in strategy with learning but is probably not worth close examination given the small differences in mean reaction time in the rankings.

The second experiment was an attempt to validate the subjective rating cards used by the Crew Performance Branch of the USAF School of Aerospace Medicine. However, these cards are of more utility when conditions are variable; e.g., subjects becoming exhausted, workloads varying from inactivity to overload, etc. In this experiment the workload presented was identical each time, and performance varied little. Some subjects reflected this in their invariate assessments. Accordingly, little correlation was expected between subjective and objective assessments and little was observed. Some subjects showed better correlations than others, particularly with workload, but the better correlations were not significantly distributed to any one of the groupings. This part of the experiment was a minor goal of the study and does not reduce the validity of the concept of subjective assessments used except, of course, in very stable conditions, as in this experiment.

Hunting for discs rather than squares in the task presented no changes in difficulty as almost identical mean reaction times were obtained for each. However, there was a considerable difference in the right-handed responses compared to left. All but two subjects were right-handed, and the mean difference undoubtedly reflects this handedness bias.

The second experiment also examined the alteration of reaction time by similar previous presentations. The memory of a previous manikin problem solved should have made the solution of a similar problem easier and hence reduce the reaction time. These analyses confirmed this hypothesis of the repetition effect.

The analysis showed that the logical similarity ranking did not correspond exactly to the reaction time ranking. Rank 2 which presented the same presentation and the same side as the previous manikin had much longer reaction times than those where side and shape were different (3 and 5). However, when all three variables were altered (types of presentation, side, and shape), the reaction times were 30 msec longer.

The results of the analysis of differential stability (Table II and Fig. 5) are encouraging. They show the repeatability of the task data from one session to another. Some instability due to learning is to be expected; note the consistently lower correlations obtained with session [illegible], but the lowest correlation observed is .56.

The plots in Figure 5 show no overall trend, up or down, the largest changes occur between sessions 10 and 11. These changes were caused by a weekend intervening for all subjects between experiments 1 and 2 of this study.

CONCLUSIONS

1. The manikin task is simple and reliable to use.

2. All subjects can learn the task and acquire plateau performance.

3. Acquisition of plateau performance is achieved in approximately [illegible] sessions of practice; the extremes range from 4 to 10 sessions.

4. The rate of acquisition of plateau performance is essentially independent of the frequency of practice or occupational level or age of the subjects.

5. The various presentations do represent varying difficulties of solution, but this variation is constant for all subjects. The sequence of presentation ensures that all 16 possibilities (and difficulties) are posed every [illegible] sec.

6. Right-handed responses were quicker than left-handed responses, but squares and discs produced similar responses.

7. There was poor correlation between subjective and objective measures of performance, but this was due to the constant control conditions used rather than the inadequacies of the performance measures.

8. A repetitive effect occurs with similarities of presentation in sequence. However, the manikin task has a fixed sequence, therefore, this effect is common for all subjects and sessions.

9. The task has a high measure of differential stability, therefore, control performance is essentially constant with time.

ACKNOWLEDGMENTS

The willing and courteous cooperation of the staff of the Crew Performance Branch of the USAF School of Aerospace Medicine is gratefully acknowledged. Special mention must be made for electronic assistance given by SMSgt F. A. Asbury, and experimental monitoring by Sgt Robin Chavez, SrA June Spinks, and A1C Vabian Paden.

REFERENCES

1. Bittner, A. C., Jr. Statistical tests for differential stability. Proceedings of the Human Factors Society 23rd Annual Meeting, Boston, Mass, 1979.

2. Denison, D. M., G. H. Byford, M. Allnutt, and D. G. Reader. Times of useful consciousness following rapid decompression from 8000 to 25,000 or 27,000 ft. 45th Symposium of Aerospace Medical Association, Washington, D.C., 1973.

3. James, W. Psychology: The Briefer Course. New York: Holt, 1893.

EN

DAT